I/O scalability boost for the next generation of Earth system models: IFS-XIOS integration as a case study

Xavier Yepes-Arbós[1], Mario C. Acosta[1], Gijs van den Oord[2], Glenn Carver[3]

[1]Barcelona Supercomputing Center (BSC), Spain, [2]Netherlands eScience Center (NLeSC), The Netherlands, [3]European Centre for Medium-Range Weather Forecasts (ECMWF), UK

AGU 100 FALL MEETING Washington, D.C. | 10–14 Dec 2018

Poster A23J-3011

1. Introduction

Overview

EC-Earth is a global coupled climate model which integrates a number of component models in order to simulate the Earth system. The two main components are **IFS as the atmospheric model** and NEMO as the ocean model.
The Integrated Forecasting System (IFS) is a global data assimilation and forecasting system developed by ECMWF. It has two different output schemes: the MF I/O server (used at ECMWF) and a **sequential I/O scheme** (used by non-ECMWF users, such as EC-Earth and OpenIFS). The sequential scheme uses an inefficient process that does not work properly when many parallel resources are used.

The I/O problem in EC-Earth

Due to this sequential scheme, the IFS version of EC-Earth has an important **I/O bottleneck**. EC-Earth was used recently to run experiments using a high resolution configuration (T511L91-ORCA025L75) under the H2020 PRIMAVERA project. Experiments required to output a lot of fields, causing a considerable slowdown in the execution time. I/O in IFS represented **about 30% of the total execution time**.

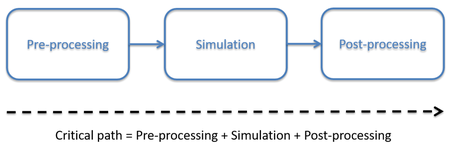
Objective

Taking advantage that NEMO already uses XIOS, this tool was chosen to be integrated into IFS as well. The XML Input/Output Server (XIOS) is an asynchronous MPI parallel I/O server developed by IPSL. The use of XIOS has **the objective of improving the computational performance and efficiency of IFS (by extension EC-Earth), and thus, reduce the execution time**.

2. The benefits of using XIOS in EC-Earth workflows

Current EC-Earth workflows

EC-Earth runs complex experiments that have different tasks in their workflows. To simplify, there are three main tasks: pre-processing, simulation and **post-processing**. They are **sequentially executed** and the time to completion is known as **critical path**. Post-processing in EC-Earth is needed to **convert GRIB files** to netCDF files, transform data to be **CMIP-compliant** (CMORization) and compute **diagnostics**. This is a very expensive process: for the tests used in this study, only the data format conversion part is 3.5x slower than the original simulation with the sequential output scheme.

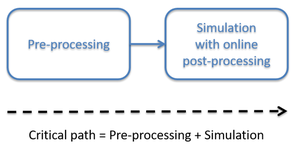


Future EC-Earth workflows

The **use of XIOS** is a key point to overcome the current post-processing issues since XIOS has the following features:

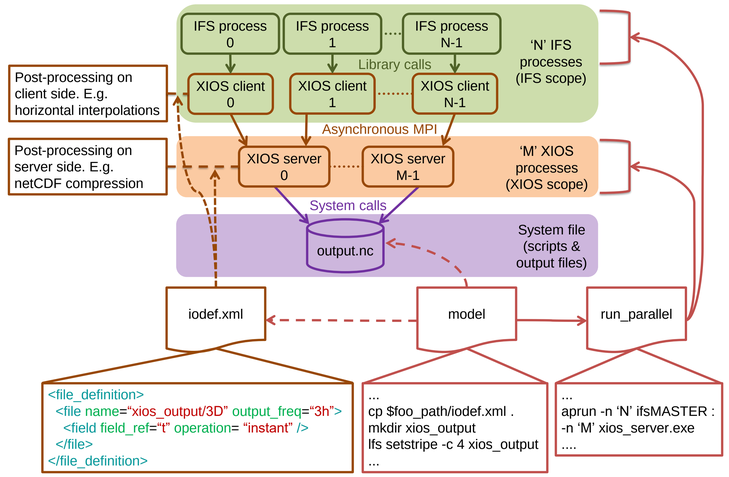
- Output files in **netCDF format**
- Output data is **CMIP-compliant** (CMORized)
- **Online post-processing** to compute diagnostics.

Critical path is shortened by **concurrently running parallel post-processing and simulation.**



3. IFS-XIOS integration

Scheme overview



FullPos integration

FullPos is a post-processing package currently used by IFS. Its use is necessary to perform **vertical interpolations** not supported by XIOS. FullPos is used before sending data from IFS to XIOS.

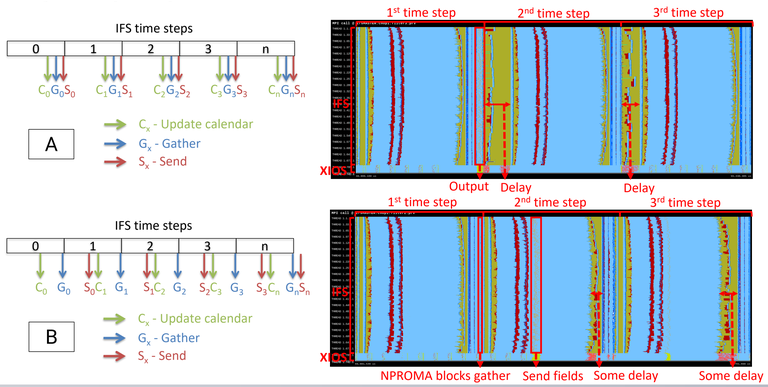
4. Performance analysis and optimization

According to the performance analysis, three different optimizations were applied:

- Parallelization using **OpenMP** threads
- **Compilation flags optimization** for external libraries and tools, e.g., I/O servers
- Computation and communication **overlap**

Computation and communication overlap

In the original scheme (A), data is sent directly from IFS to XIOS. The result is that there was not an appropriate overlap between IFS computation and communications to XIOS. In the new scheme (B), communications were delayed to **truly overlap computation and communication**.

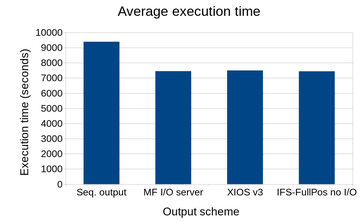


5. Performance evaluation

Execution environment

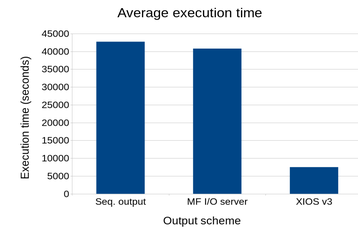
ECMWF HPC platform (Cray XC40); IFS CY43R3; Octahedral reduced Gaussian grid: T1279 (16 km); 702 MPI x 6 OpenMP; 10 days forecast with a time step of 600 seconds; 3-hourly output files; only grid-point fields; netCDF files size: 2.5 TB.

Comparison test between different output schemes



- Sequential output scheme: 9391 seconds (20.7% of overhead)
- IFS-XIOS non-optimized: 7682 seconds (**3.1% of overhead**)
- IFS-XIOS optimized: 7499 seconds (**0.7% of overhead**)
- XIOS is as fast as the MF I/O server and comparable to no I/O.
- In **less than 1 minute** of overhead **IFS outputs 2.5 TB** of data

Comparison test adding GRIB to netCDF post-processing



- The post-processing takes 9.2 hours (sequentially performed, as in EC-Earth)
- IFS-XIOS optimized is **5.7x faster** than the sequential output
- IFS-XIOS optimized is **5.4x faster** than the MF I/O server

6. Conclusions and future work

Conclusions

- We present a new **easy-to-use** I/O scheme for IFS using a parallel I/O server called XIOS. The development is able to output data using the netCDF format with almost no overhead (**only 0.7%**). This avoids the format conversion from GRIB to netCDF, being the new integration **5.7x faster** than the current output approach used by the EC-Earth community
- The integration with no optimizations already improved the execution time
- The integration with optimizations is scalable, fast, efficient and will **address the I/O issue**
- Summary of **benefits for EC-Earth 4**:
  - Increase the performance and efficiency of the model
  - Online diagnostics computation
  - CMORized netCDF files
  - Data compression
  - Simpler output configuration file using XML syntax
  - Experiments with simpler workflows
  - Save thousands of computing hours and storage space

Future work

- Transform fields from spectral space to grid-point space and send them to XIOS
- The development done for IFS will be **ported to OpenIFS** (next release)
- Adapt the future EC-Earth 4 version to output fields and compute online diagnostics from OpenIFS and NEMO components through XIOS

esiwace CENTRE OF EXCELLENCE IN SIMULATION OF WEATHER AND CLIMATE IN EUROPE

The research leading to these results has received funding from the EU H2020 Framework Programme under grant agreement no. 675191 (ESiWACE).
This material reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.
This poster can be downloaded at **https://earth.bsc.es/wiki/doku.php?id=library:external:posters**
The report can be downloaded at **https://earth.bsc.es/wiki/doku.php?id=library:external:technical_memoranda**
Corresponding author: **xavier.yepes@bsc.es**

DOWNLOAD THE POSTER

DOWNLOAD THE REPORT